

#2

05-90
0228

OIPE

RAW SEQUENCE LISTING

DATE: 03/21/2002

PATENT APPLICATION: US/10/071,751

TIME: 15:52:05

Input Set : N:\Crf3\RULE60\10071751.txt

Output Set: N:\CRF3\03212002\J071751.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Hunter, Shirley Wu
 6 Sim, Gek-Kee
 7 Weber, Eric R.

9 (ii) TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
 10 APPARATUS TO COLLECT SUCH PROTEINS

12 (iii) NUMBER OF SEQUENCES: 88

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: SHERIDAN ROSS P.C.
 16 (B) STREET: 1560 BROADWAY, SUITE 1200
 17 (C) CITY: DENVER
 18 (D) STATE: CO
 19 (E) COUNTRY: U.S.A.
 20 (F) ZIP: 80202

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk
 24 (B) COMPUTER: IBM PC compatible
 25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/10/071,751
 C--> 30 (B) FILING DATE: 07-Feb-2002
 31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 09/171,156
 35 (B) FILING DATE: 1998-10-09

39 (viii) ATTORNEY/AGENT INFORMATION:

40 (A) NAME: Connell, Gary J.
 41 (B) REGISTRATION NUMBER: 32,020
 42 (C) REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: 303/863-9700
 46 (B) TELEFAX: 303/863-0223

49 (2) INFORMATION FOR SEQ ID NO: 1:

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 26 amino acids
 53 (B) TYPE: amino acid
 54 (C) STRANDEDNESS:
 55 (D) TOPOLOGY: linear

57 (ii) MOLECULE TYPE: protein

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 Met Arg Gly Asn His Val Phe Leu Glu Asp Gly Met Ala Asp Met Thr

ENTERED

RAW SEQUENCE LISTING

DATE: 03/21/2002

PATENT APPLICATION: US/10/071,751

TIME: 15:52:05

Input Set : N:\Crf3\RULE60\10071751.txt

Output Set: N:\CRF3\03212002\J071751.raw

```

62          1          5          10          15
64      Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr
65          20          25
67 (2) INFORMATION FOR SEQ ID NO: 2:
68     (i) SEQUENCE CHARACTERISTICS:
69         (A) LENGTH: 12 amino acids
70         (B) TYPE: amino acid
71         (C) STRANDEDNESS:
72         (D) TOPOLOGY: linear
73     (ii) MOLECULE TYPE: protein
74     (ix) FEATURE:
75         (A) NAME/KEY: Xaa = Tyr or Asp
76         (B) LOCATION: 5
77     (ix) FEATURE:
78         (A) NAME/KEY: Xaa = any amino acid
79         (B) LOCATION: 6
80     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
W--> 88      Lys Tyr Arg Asn Xaa Xaa Thr Asn Asp Pro Gln Tyr
89          1          5          10
91 (2) INFORMATION FOR SEQ ID NO: 3:
92     (i) SEQUENCE CHARACTERISTICS:
93         (A) LENGTH: 27 amino acids
94         (B) TYPE: amino acid
95         (C) STRANDEDNESS:
96         (D) TOPOLOGY: linear
97     (ii) MOLECULE TYPE: protein
98     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
101      Glu Ile Lys Arg Asn Asp Arg Glu Pro Gly Asn Leu Ser Lys Ile Arg
102          1          5          10          15
103      Thr Val Met Asp Lys Val Ile Lys Gln Thr Gln
104          20          25
106 (2) INFORMATION FOR SEQ ID NO: 4:
107     (i) SEQUENCE CHARACTERISTICS:
108         (A) LENGTH: 23 amino acids
109         (B) TYPE: amino acid
110         (C) STRANDEDNESS:
111         (D) TOPOLOGY: linear
112     (ii) MOLECULE TYPE: protein
113     (ix) FEATURE:
114         (A) NAME/KEY: Xaa = Ala or His
115         (B) LOCATION: 8
116     (ix) FEATURE:
117         (A) NAME/KEY: Xaa = Ala or His
118         (B) LOCATION: 9
119     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
W--> 129      Leu Lys Asp Asn Asp Ile Tyr Xaa Xaa Arg Asp Ile Asn Glu Ile Leu
130          1          5          10          15
131      Arg Val Leu Asp Pro Ser Lys
132          20
133

```

RAW SEQUENCE LISTING

DATE: 03/21/2002

PATENT APPLICATION: US/10/071,751

TIME: 15:52:05

Input Set : N:\Crf3\RULE60\10071751.txt

Output Set: N:\CRF3\03212002\J071751.raw

135 (2) INFORMATION FOR SEQ ID NO: 5:

137 (i) SEQUENCE CHARACTERISTICS:

138 (A) LENGTH: 27 amino acids

139 (B) TYPE: amino acid

140 (C) STRANDEDNESS:

141 (D) TOPOLOGY: linear

143 (ii) MOLECULE TYPE: protein

145 (ix) FEATURE:

146 (A) NAME/KEY: Xaa = any amino acid

147 (B) LOCATION: 12

149 (ix) FEATURE:

150 (A) NAME/KEY: Xaa = any amino acid

151 (B) LOCATION: 18

153 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

W--> 155 Asn Tyr Gly Arg Val Gln Ile Glu Asp Tyr Thr Xaa Ser Asn His Lys

156 1 5 10 15

W--> 158 Asp Xaa Glu Glu Lys Asp Gln Ile Asn Gly Leu

159 20 25

161 (2) INFORMATION FOR SEQ ID NO: 6:

163 (i) SEQUENCE CHARACTERISTICS:

164 (A) LENGTH: 18 amino acids

165 (B) TYPE: amino acid

166 (C) STRANDEDNESS:

167 (D) TOPOLOGY: linear

169 (ii) MOLECULE TYPE: protein

171 (ix) FEATURE:

172 (A) NAME/KEY: Xaa = any amino acid

173 (B) LOCATION: 5

176 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

W--> 178 Lys Tyr Arg Asn Xaa Tyr Thr Asn Asp Pro Gln Leu Lys Leu Leu Asp

179 1 5 10 15

181 Glu Gly

184 (2) INFORMATION FOR SEQ ID NO: 7:

186 (i) SEQUENCE CHARACTERISTICS:

187 (A) LENGTH: 22 amino acids

188 (B) TYPE: amino acid

189 (C) STRANDEDNESS:

190 (D) TOPOLOGY: linear

192 (ii) MOLECULE TYPE: protein

194 (ix) FEATURE:

195 (A) NAME/KEY: Xaa = any amino acid

196 (B) LOCATION: 13

198 (ix) FEATURE:

199 (A) NAME/KEY: Xaa = any amino acid

200 (B) LOCATION: 19

202 (ix) FEATURE:

203 (A) NAME/KEY: Xaa = any amino acid

204 (B) LOCATION: 21

206 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

RAW SEQUENCE LISTING

DATE: 03/21/2002

PATENT APPLICATION: US/10/071,751

TIME: 15:52:05

Input Set : N:\Crf3\RULE60\10071751.txt

Output Set: N:\CRF3\03212002\J071751.raw

```

W--> 208      Tyr Phe Asn Asp Gln Ile Lys Ser Val Met Glu Pro Xaa Val Phe Lys
      209      1              5              10              15
W--> 211      Tyr Pro Xaa Ala Xaa Leu
      212      20
214 (2) INFORMATION FOR SEQ ID NO: 8:
216 (i) SEQUENCE CHARACTERISTICS:
217 (A) LENGTH: 20 base pairs
218 (B) TYPE: nucleic acid
219 (C) STRANDEDNESS: single
220 (D) TOPOLOGY: linear
222 (ii) MOLECULE TYPE: DNA (genomic)
224 (ix) FEATURE:
225 (A) NAME/KEY: misc_feature
226 (B) LOCATION: 1..20
227 (D) OTHER INFORMATION: /label= primer
229 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
231 TGRITTCWA TRAARTCTTC 20
233 (2) INFORMATION FOR SEQ ID NO: 9:
235 (i) SEQUENCE CHARACTERISTICS:
236 (A) LENGTH: 225 base pairs
237 (B) TYPE: nucleic acid
238 (C) STRANDEDNESS: single
239 (D) TOPOLOGY: linear
241 (ii) MOLECULE TYPE: DNA (genomic)
243 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
245 GAATTCGGCA CGAGTGAAAT TCAATATTTT GTTTTACATT AAATTTTTC AATTCGATAT 60
247 GAAATTTTTA CTGGCAATTT GCGTGTGTG TGTTTATTA AATCAAGTAT CTATGTCAAA 120
249 AATGGTCACT GAAAAGTGTA AGTCAGGTGG AAATAATCCA AGTACAGAAG AGGTGTCAAT 180
251 ACCATCTGGG AAGCTTACTA TTGAAGATTT TTGTATTGGA AATCA 225
253 (2) INFORMATION FOR SEQ ID NO: 10:
255 (i) SEQUENCE CHARACTERISTICS:
256 (A) LENGTH: 15 base pairs
257 (B) TYPE: nucleic acid
258 (C) STRANDEDNESS: single
259 (D) TOPOLOGY: linear
261 (ii) MOLECULE TYPE: DNA (genomic)
263 (ix) FEATURE:
264 (A) NAME/KEY: misc_feature
265 (B) LOCATION: 1..15
266 (D) OTHER INFORMATION: /label= primer
268 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
270 AATTCGGCAC GAGTG 15
272 (2) INFORMATION FOR SEQ ID NO: 11:
274 (i) SEQUENCE CHARACTERISTICS:
275 (A) LENGTH: 565 base pairs
276 (B) TYPE: nucleic acid
277 (C) STRANDEDNESS: single
278 (D) TOPOLOGY: linear
280 (ii) MOLECULE TYPE: cDNA

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/071,751

DATE: 03/21/2002

TIME: 15:52:05

Input Set : N:\Crf3\RULE60\10071751.txt

Output Set: N:\CRF3\03212002\J071751.raw

```

282 (ix) FEATURE:
283 (A) NAME/KEY: CDS
284 (B) LOCATION: 45..314
287 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
289 TGAAATTCAA TATTTTGTTT TACATTAAAT TTTTCAAATT CGAT ATG AAA TTT TTA 56
290 Met Lys Phe Leu
291 1
293 CTG GCA ATT TGC GTG TTG TGT GTT TTA TTA AAT CAA GTA TCT ATG TCA 104
294 Leu Ala Ile Cys Val Leu Cys Val Leu Leu Asn Gln Val Ser Met Ser
295 5 10 15 20
297 AAA ATG GTC ACT GAA AAG TGT AAG TCA GGT GGA AAT AAT CCA AGT ACA 152
298 Lys Met Val Thr Glu Lys Cys Lys Ser Gly Gly Asn Asn Pro Ser Thr
299 25 30 35
301 GAA GAG GTG TCA ATA CCA TCT GGG AAG CTT ACT ATT GAA GAT TTT TGT 200
302 Glu Glu Val Ser Ile Pro Ser Gly Lys Leu Thr Ile Glu Asp Phe Cys
303 40 45 50
305 ATT GGA AAT CAT CAA AGT TGC AAA ATA TTT TAC AAA AGT CAA TGT GGA 248
306 Ile Gly Asn His Gln Ser Cys Lys Ile Phe Tyr Lys Ser Gln Cys Gly
307 55 60 65
309 TTT GGA GGT GGT GCT TGT GGA AAC GGT GGT TCA ACA CGA CCA AAT CAA 296
310 Phe Gly Gly Gly Ala Cys Gly Asn Gly Gly Ser Thr Arg Pro Asn Gln
311 70 75 80
313 AAA CAC TGT TAT TGC GAA TAACCATATT CCGGATGAAA GACCAAATTG 344
314 Lys His Cys Tyr Cys Glu
315 85 90
317 ATATAAATTA CTAATAATTAT GCTAGATAGC AATCATAAAA TTTTGAAGTT TTCAATGATC 404
319 CTAACATGTT TTGCCTCCAA TTTATTTTAA CAGCAAATTG CTGGAACCTTA CCGTACCGTA 464
321 ACTAAATGTT CAAGAAATAC TGAATGTTTA CAAATAGATT ATTATAAATA TTGTAACATT 524
323 GTCTAATATT TATAAGAATT ATATAAACTG AATTGCAAAA A 565
325 (2) INFORMATION FOR SEQ ID NO: 12:
327 (i) SEQUENCE CHARACTERISTICS:
328 (A) LENGTH: 90 amino acids
329 (B) TYPE: amino acid
330 (D) TOPOLOGY: linear
332 (ii) MOLECULE TYPE: protein
334 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
336 Met Lys Phe Leu Leu Ala Ile Cys Val Leu Cys Val Leu Leu Asn Gln
337 1 5 10 15
339 Val Ser Met Ser Lys Met Val Thr Glu Lys Cys Lys Ser Gly Gly Asn
340 20 25 30
342 Asn Pro Ser Thr Glu Glu Val Ser Ile Pro Ser Gly Lys Leu Thr Ile
343 35 40 45
345 Glu Asp Phe Cys Ile Gly Asn His Gln Ser Cys Lys Ile Phe Tyr Lys
346 50 55 60
348 Ser Gln Cys Gly Phe Gly Gly Gly Ala Cys Gly Asn Gly Gly Ser Thr
349 65 70 75 80
351 Arg Pro Asn Gln Lys His Cys Tyr Cys Glu
352 85 90
354 (2) INFORMATION FOR SEQ ID NO: 13:

```

RAW SEQUENCE LISTING ERROR SUMMARY
 PATENT APPLICATION: US/10/071,751

DATE: 03/21/2002
 TIME: 15:52:06

Input Set : N:\Crf3\RULE60\10071751.txt
 Output Set: N:\CRF3\03212002\J071751.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; Xaa Pos.5,6
 Seq#:4; Xaa Pos.8,9
 Seq#:5; Xaa Pos.12,18
 Seq#:6; Xaa Pos.5
 Seq#:7; Xaa Pos.13,19,21
 Seq#:29; Xaa Pos.379
 Seq#:30; Xaa Pos.379
 Seq#:35; N Pos. 192
 Seq#:70; Xaa Pos.1,24

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:8; Line(s) 231
 Seq#:9; Line(s) 245,247,249,251
 Seq#:10; Line(s) 270
 Seq#:11; Line(s) 289,293,297,301,305,309,313,317,319,321,323
 Seq#:13; Line(s) 370,374,378,382,386,390
 Seq#:15; Line(s) 440
 Seq#:16; Line(s) 459
 Seq#:17; Line(s) 477,479,483,487,491,495,500,504,508,512,516,520,524,526
 Seq#:17; Line(s) 528,530,532
 Seq#:19; Line(s) 587,589,591,593,595,597,599,601
 Seq#:20; Line(s) 619,623,627,631,635,639,643,647,651,655,659,663,667,671
 Seq#:20; Line(s) 675,679,683,687,691,695,699,703,707,711,715,719,723,727
 Seq#:20; Line(s) 731,735,739,743,747,751,755,759,763,767,771,775,779,783
 Seq#:20; Line(s) 787,791,795,799,803,807,811,815,819,823,827,831,835,839
 Seq#:20; Line(s) 843
 Seq#:22; Line(s) 1045,1049,1053,1057,1061,1065,1070,1074,1078
 Seq#:24; Line(s) 1136,1140,1144,1148,1152,1156
 Seq#:26; Line(s) 1205,1209,1213,1217,1221,1225,1229,1233,1237,1241,1245
 Seq#:26; Line(s) 1249,1254,1258,1262,1266,1270,1274,1278,1282,1286,1290
 Seq#:26; Line(s) 1294,1298,1302,1306,1310,1314,1318,1322,1326,1328,1330
 Seq#:26; Line(s) 1332,1334
 Seq#:28; Line(s) 1446,1448,1450,1452,1454,1456,1458,1460,1462,1464,1466
 Seq#:28; Line(s) 1468,1470,1472,1474,1476,1478,1480,1482,1484,1486,1488
 Seq#:28; Line(s) 1490,1492
 Seq#:29; Line(s) 1514,1518,1522,1526,1530,1534,1538,1542,1546,1550,1554
 Seq#:29; Line(s) 1558,1562,1566,1570,1574,1578,1582,1586,1590,1594,1598
 Seq#:29; Line(s) 1602,1606,1610,1614,1618,1622,1626,1630,1634,1638,1642
 Seq#:29; Line(s) 1646,1650,1654,1658
 Seq#:31; Line(s) 1801,1803,1805,1807,1809
 Seq#:32; Line(s) 1823,1825,1827,1829,1831,1833

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/071,751DATE: 03/21/2002
TIME: 15:52:06Input Set : N:\Crf3\RULE60\10071751.txt
Output Set: N:\CRF3\03212002\J071751.raw

Seq#:33; Line(s) 1847,1849,1851,1853,1855,1857,1859
Seq#:34; Line(s) 1873,1875,1877,1879,1881
Seq#:35; Line(s) 1895,1897,1899,1901
Seq#:36; Line(s) 1919,1923,1927,1931,1935,1939,1943,1947
Seq#:38; Line(s) 2001,2005,2009,2013,2017,2021,2025,2029,2033,2037
Seq#:40; Line(s) 2089,2091,2093,2095,2097,2099
Seq#:41; Line(s) 2113,2115,2117,2119,2121,2123,2125,2127,2129
Seq#:42; Line(s) 2143,2145,2147,2149,2151,2153,2155,2157,2159,2161,2163
Seq#:43; Line(s) 2177,2179,2181,2183,2185,2187,2189,2191
Seq#:44; Line(s) 2205,2207,2209,2211,2213
Seq#:45; Line(s) 2227,2229,2231,2233,2235,2237,2239
Seq#:46; Line(s) 2253,2255,2257,2259,2261
Seq#:47; Line(s) 2275,2277,2279,2281,2283,2285,2287,2289,2291,2293,2295
Seq#:48; Line(s) 2313,2317,2321,2325,2329,2333,2337,2341,2345,2349,2353
Seq#:48; Line(s) 2357
Seq#:50; Line(s) 2424,2428,2432,2436,2440,2444,2448,2452,2456
Seq#:52; Line(s) 2514,2518,2522,2526,2530,2534,2538,2542,2544,2546,2548
Seq#:52; Line(s) 2550
Seq#:54; Line(s) 2593,2595,2597,2599,2601,2603,2605,2607,2609,2611
Seq#:55; Line(s) 2629,2633,2637,2641,2645,2649
Seq#:57; Line(s) 2694,2696,2698,2700,2702
Seq#:58; Line(s) 2720,2724,2728,2732,2736
Seq#:60; Line(s) 2779,2781,2783,2785
Seq#:61; Line(s) 2803,2807,2811,2815,2819,2823,2827,2831,2835,2839,2843
Seq#:61; Line(s) 2845,2847,2849,2851,2853,2855,2857,2859
Seq#:63; Line(s) 2914,2916,2918,2920,2922,2924,2926,2928,2930,2932,2934
Seq#:63; Line(s) 2936,2938,2940,2942,2944,2946
Seq#:64; Line(s) 2964,2968,2972,2976,2980,2984,2988,2992,2996,3000,3004
Seq#:64; Line(s) 3008,3012,3016,3020,3024,3028,3032,3036,3040,3044,3048
Seq#:64; Line(s) 3052,3056,3058
Seq#:66; Line(s) 3152,3154,3156,3158,3160,3162,3164,3166,3168,3170,3172
Seq#:66; Line(s) 3174,3176,3178,3180,3182,3184,3186,3188,3190,3192
Seq#:67; Line(s) 3211,3215,3219,3223,3227,3231,3235,3239,3243,3247,3251
Seq#:67; Line(s) 3255,3259,3263,3267,3271,3275,3279,3283,3287,3291,3295
Seq#:67; Line(s) 3299
Seq#:69; Line(s) 3395,3397,3399,3401,3403,3405,3407,3409,3411,3413,3415
Seq#:69; Line(s) 3417,3419,3421,3423,3425,3427,3429
Seq#:71; Line(s) 3475,3479,3483,3487,3491,3495,3499,3503,3507
Seq#:73; Line(s) 3562,3564,3566,3568,3570,3572,3574
Seq#:74; Line(s) 3593,3597,3601,3605,3609,3613,3615,3617
Seq#:76; Line(s) 3657,3659,3661,3663,3665,3667,3669
Seq#:79; Line(s) 3733
Seq#:80; Line(s) 3752
Seq#:81; Line(s) 3771
Seq#:82; Line(s) 3790
Seq#:83; Line(s) 3809
Seq#:84; Line(s) 3828
Seq#:85; Line(s) 3847
Seq#:86; Line(s) 3866

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/071,751

DATE: 03/21/2002
TIME: 15:52:06

Input Set : N:\Crf3\RULE60\10071751.txt
Output Set: N:\CRF3\03212002\J071751.raw

Seq#:88; Line(s) 3907

VERIFICATION SUMMARY

DATE: 03/21/2002

PATENT APPLICATION: US/10/071,751

TIME: 15:52:06

Input Set : N:\Crf3\RULE60\10071751.txt

Output Set: N:\CRF3\03212002\J071751.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:88 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:129 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:155 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:158 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:16
L:178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:208 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
L:211 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:16
L:1607 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:1152
L:1747 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:368
L:3452 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70 after pos.:0
L:3455 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70 after pos.:16